MPEG 비디오 색인과 검색을 위한 색인키 생성 알고리즘

하성욱[†] · 강대성[††] · 김대진[†††] · 권기항[††††]

요 약

본 논문에서는 MPEG 비디오 데이터들을 효율적으로 색인하고 검색하기 위한 키를 정의하는 새로운 방식을 제안한다. 기존의 방법들은 추출된 Shot과 사용자의 질의 영상을 여러 가지 특징으로 패턴 매칭하므로 수행시간이 오래 걸린다. 이에 반해, 제안하는 PKey 생성 방법은 추출된 PFrame의 영상을 여러 가지 특징으로 영상을 특정한 비트로 코드화하여 저장하므로 색인과 검색시에 사용자가 질의한 영상의 PKey를 사용하여 수행속도를 높일 수 있다. 사용자 질의에 대한 투명성을 위해 비구문 기반 서술에 대한 5가지의 특징을 PKey로서 정의하였다.

An algorithm for creating index key for MPEG videos

Seong-Wook Ha[†], Dae-Seong Kang[††], Daijin Kim[†††] and Keehang Kwon[††††]

ABSTRACT

In this paper, we suggest a novel method which defines index for effectively searching MPEG video data. The traditional methods are slow because they rely on pattern matching of images between extracted shots and user's querying images. Whereas, our method uses a different approach: It transforms images of extracted PFrames into codes and uses the codes for creating index keys. To be specific, we define five features of non-context description for transparency about user's querying.

1. 서 론

MPEG의 새로운 흐름으로서 압축과 복원뿐만 아니라, 현재의 멀티미디어 데이터들을 일괄적으로 관리할 수 있는 방법으로 MPEG-7[1] 표준을 제정하기 위한 연구가 활발히 진행중이며, 멀티미디어 데이터베이스에서 데이터를 효율적으로 표현하는 멀티미디어 서술자에 대해서 연구되고 있다. 멀티미디어 서술자(descriptor)는 크게 구문 기반 서술자와 비구문 기반 서술자로 나눌 수 있다. 구문 기반 서술자에 비해 비구문 기반 서술자는 멀티미디어 파일에 대한 관리자 측면의 서술이 필요 없으며, 견해에 따른 응답 오류를 줄일 수 있다. 본 논문에서는 비구문 기반 서술자를 위하여 새롭게 생성된 MPEG 기반의 멀티미디어 파일과 기존의 멀티미디어 데이터베이스에서 색인이나 검색에서 필요한 키(key)가 되는 키 프레임(key frame)을 선택하는 방법에 대하여 제안할 것이다. 키 프레임 생성 전에 큰 특징 변화시 대표 프레임이 검출되며, 여러 가지 특징 프레임에서 칼라, 모양, 위치 등의 특징에서 비슷한 특징을 갖는 프레임을 통합하여 정보를 줄일 것이며, 후보 영상의 수를 줄여 속도를 높일 것이다. 대표 프레임들을 선택하며 여기서 PFrame(primary frame)을 재선택하여야 한다. 본 연구에서는 PFrame에 대한 정의와 특징 프레임의 개수를 줄이는 방법과 PKey 추출에 대

이 논문은 1998학년도 동아대학교 학술연구조성비(특정목적)에 의하여 연구되었음

[†] 동아대학교 컴퓨터공학과(박사과정 재학중)
[††] 동아대학교 전기전자컴퓨터공학부 조교수
[†††] 포항공과대학교 컴퓨터공학과 부교수
[††††] 동아대학교 컴퓨터공학과 조교수

한 방법을 제안하고자 한다. 우선, 전체 프레임에서 m구간내에서 칼라 히스토그램이 크게 변화되는 프레임이 특징 프레임으로 정의되며, 프레임들은 여러 가지 특징중 색상에 대하여 서로 비슷한 특징을 갖게 된다. 추출된 특징 프레임들은 시간 순서상 특징의 상관이 높은 프레임이 존재하게 되므로, 전체 특징 프레임들에서 다시 색상에 대한 기준 문턱치를 조절하여 특징 프레임들의 수를 줄였다. 재추출된 특징 프레임들을 PFrame으로 정의한다. 추출된 PFrame에서 색인과 검색을 위한 키 정의가 필요하며, 효율적인 키를 구축하는 색인과 검색에 투명성을 부여할 수 있다. 본 연구에서는 이러한 키를 PKey로서 정의하며, 주 색상 · 방향 · 직선 및 곡선 · PCA 정보를 추출하여 저장한다. 주 색상 정보는 해당 색상의 히스토그램에서 피크치에 대한 퍼지 합집합을 이용한 양자화에서 추출되는 주 색상 성분을 얻으며, 방향 정보는 퍼지 벌레 검색을 사용하여 에지 검출에서 정의하는 방향 매스크를 사용하는 방법으로 방향성을 얻게 된다. 직선 및 곡선 정보는 퍼지 벌레 검색에서 매스크의 인덱스 정보로서 연결되는 직선과 곡선 개수를 정보로 저장하며, PCA 정보는 해당 색상에 대한 노드 분할과 주축 각도 계산 알고리즘을 이용하여 각 색상에 대한 주 성분 정보를 추출하게 된다. 추출된 4가지 정보를 규합하여 하나의 키로서 저장하게 된다. 이로서, 전체 영상을 검색하지 않고 정의된 PKey만으로 색인과 검색에 사용할 수 있을 것이다. 제안한 알고리즘은 질의에 대한 정규화된 응답을 하는 비구문 기반 서술자를 만들기 위한 기반 기술이며, 특징점들을 PKey로 정의하여 단순화하므로 데이터베이스 공간 활용을 높일 수 있다.

2. 색인 서술자

멀티미디어 정보에 대한 검색을 하기 위해서는 일정한 규칙을 두고 데이터들에 대해 색인을 해야 하며, 사용자는 이러한 규칙에 해당하는 방법으로 데이터들을 검색할 수가 있다[2]. 멀티미디어 데이터베이스에 저장된 정보에 대한 검색과 색인을 위해 사용되는 서술자는 크게 두 가지로 나눌 수 있다[3]. 첫 번째 서술자는 미디어에서 정보의 특성을 구문으로 서술하는 구문 기반 서술자이다. 구문 기반 서술자는 미디어 데이터를 올바르게 표현할 수 있는 문장이나 단어로 나타내는 방식이다. 두 번째 서술자는 멀티미디어 정보가 가지는 특징을 일관된 표현 방식으로 나타내는 비구문 기반 서술자[4]이다. 전자의 방식은 표현과 구현이 쉬운 반면에, 색인된 데이터를 검색할 경우, 검색하는 사용자와 데이터를 표현한 오퍼레이터간의 구문 표현 능력의 차이로 검색이 제대로 이루어지지 않을 수 있으며, 데이터의 특징을 표현할 때 개인적인 차이로 정확히 데이터를 표현하지 못 할 수도 있다. 또한, 대용량 정보의 표현은 많은 시간과 인력이 요구된다. 이에 반해, 후자는 여러 데이터에 대하여 일관적인 방식으로 정의하므로, 색인을 일반화하며 검색에서도 자체 표준화한 양식에 따르게 되어 개인별 오차를 없앨 수 있다. 두 방식에 대하여 표. 1에서 비교하였다. 멀티미디어 데이터의 색인은 멀티미디어 데이터가 어떻게 표현되는가, 사용자가 어떤 검색을 하고자 하는가에 좌우된다. 색인시 기본적으로 고려해야할 사항은 하나의 멀티미디어 내에서 여러 구간에 대해서 중첩된 색인을 할 수 있어야 한다.

표 1. 구문 기반 서술자와 비구문 기반 서술자의 비교

	구문 기반 서술자	비구문 기반 서술자
특징 정의	쉽다.	어렵다.
표현 방식	일반화되지 않는다.	자체 서술자에 따르는 곳에서는 일반화된다.
색인 방식	개인 오차가 심하다.	일관적으로 표시한다.
검색 결과	개인 오차로 제대로 이루어지지 않는다.	정의한 특징에 따라 달라진다.
수행 속도	구문 길이에 비례한다.	특징 개수에 비례한다.

3. PFrame 추출

본 연구에서는 기존의 규칙에 의존하지 않고 사용자가 독립적으로 검색할 수 있는 비구문 기반 서술자를 위한 특징 추출이 목적이다. 멀티미디어 정보에서 비구문 기반 서술자를 위한 프레임을 추출하기 위해서는 멀티미디어 정보 중에서 가장 주요한 특징 프레임을 추출하는 것이 중요하다. 여기서, 주요 프레임을 PFrame이라 칭하며, MPEG-1과 MPEG-2로 압축된 비디오에서 가장 두드러지는 프레임으로 정의된다. MPEG-1은 미디어 저장을 위해서 표준화된 압

축 형식이며, MPEG-2[5]는 그에 비해 디지털 방송을 위해 정의된 압축 형식으로 더욱 복잡한 알고리즘을 갖고 있다. 우선, MPEG-1과 MPEG-2에서 추출할 정보를 얻기 위해서는 압축된 영역에서 비디오에 해당하는 부분을 복원하는 과정이 필요하다. 복원된 비디오 영역에서는 이전 프레임과 현재 프레임의 스트림(stream) 상에서 변화가 큰 영역을 추출할 수 있는 특징점 정의가 필요하다. 비디오의 스트림에서 시간당 칼라 히스토그램의 변화를 조사하여, 칼라 히스토그램이 급격히 변화하는 프레임을 우선 대표 프레임으로 검출한다[6]. 이렇게 검출된 프레임은 비디오, 장면 변화, 카메라의 흔들림, 조광 변화, 프로그0램(전체 비디오의 구성)의 변화 등으로 판단할 수 있다. 이러한 변화 시에 검출된 프레임들은 한 비디오에 대하여 표현될 수 있는 데이터로는 충분하나, 똑같은 특징들이 여러장 반복될 수 있고 Fade In/Out 효과, 오버랩 효과, Z축-이동, 줌 In/Out 효과 등에는 적용되지 못한다[7]. 점진적으로 변화하는 히스토그램에 대한 해결책으로 증감되는 변화량의 절대치 합을 이용할 수 있다. 절대치 합에 대하여 비디오의 크기, 해상도(resolution)가 문턱치 이상이면 그 시점을 다른 대표 프레임으로 추출한다[8]. 비디오의 장면 변화로 볼 수 있는 것은 색상의 변화와 새로운 객체의 출현으로 볼 수 있다. 아래의 알고리즘에서는 새로 생성된 PFrame과 현재까지의 PFrame들을 비교하여 상관이 높은 PFrame이 존재하는 경우 생성된 PFrame을 제거하여, 시간적으로 동떨어진 환경에서 시스템의 메모리 활용도를 높인다. 대표 프레임 추출 알고리즘은 다음과 같다.

3.1 대표 프레임 추출 알고리즘

영상의 대표 프레임은 칼라 히스토그램 변화를 검출한다. MPEG은 높은 압축률을 위해 YUV의 좌표계를 가지고 있다. 본 연구에서는 컴퓨터 환경상에서 주로 표현되는 RGB의 좌표계로 변환 후에, R · G · B 히스토그램 각각의 변화량과 전체 히스토그램의 변화량을 우선 대표 프레임을 추출하는 데에 사용한다. 우선, R · G · B 레벨 히스토그램을 각 프레임마다 계산한다. 계산된 히스토그램에서, m 구간 전후의 프레임에서 일정한 문턱치 이상이 되면 대표 프레임으로 검출한다. m 구간은 장면 변화가 서서히 일어나는 Fade In, Fade Out, Zoom, Morphing과 같은 변화를 적절히 검출할 수 있는 구간으로 정의한다. 2×m 구간에서 각 히스토그램의 변화를 R_{Total}, G_{Total}, B_{Total}에 저장한다. 3가지 변수에서 일정 문턱치 이상이면, 2×m 구간에서 장면 변화가 발생한 것으로 간주하고 구간내에서 분산값이 가장 큰 장면을 대표 프레임으로 추출한다. 히스토그램에 대한 문턱치는 3가지 좌표 모두 동일한 값을 사용하여 검출한다. 비디오의 특성에 따라서, 각 히스토그램에서 대표 프레임이 추출되는 빈도는 상이하다. 두 번째 대표 프레임 검출 조건은 프레임별 히스토그램에서 공간 정보의 상관성으로 검출한다. 히스토그램 공간 정보는 히스토그램 레벨을 전체 프레임 크기의 확률로 정의한다. 정의된 공간의 블록들은 2×m 구간에서 검출 후, 공간상의 블록 변화가 크게 나타나는 프레임을 대표 프레임으로 추출한다. b 크기에 의해서 설정되는 블록들은 히스토그램의 확률로 표현되는 크기이다. 그림 1은 히스토그램에서 나타나는 분포를 확률 블록 b로 정의하여 표현한 그림이다.

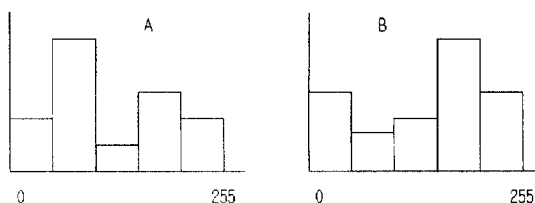



그림 1. 히스토그램의 공간 분포

그림 1에서 A와 B는 각 프레임에서 표현되는 히스토그램 분포를 확률로 표현한 그래프이다. 전후 프레임에 대한 히스토그램의 공간적인 차분이 발생될 수 있으며, 이러한 변화량으로서 장면 변화를 검출할 수 있다. 블록 크기 b에 따라, 장면 변화의 검출이 달라진다. 블록의 개수를 임의의 bs로 정의할 때 블록 크기 b는 다음과 같이 정의된다.

$$b = \frac{Frame_{Height} * Frame_{Width}}{bs}$$

또한, b는 히스토그램에 대한 확률적인 분포에 대한 정보로서, 아래와 같이 정의된다.

$$b_k = P(Frame_{Hist,} | Frame_{Size})$$

따라서, 전체 프레임 크기에 대한 히스토그램의 i 번째 레벨의 확률로서 표현된다. 공간 정보는 R · G · B 히스토그램에 대한 공간 정보로서 추출되며, 2×m 구간에서의 변화량을 조사하여 서서히 변화하는 프레임에 대한 검출도 가능하다. 각 좌표의 공간 정보는 공간 블록 크기 b에 의존하며, 공간 상관성에 따라 적절한 블록 크기로서 장면 변화를 검출한다. 이전 프레임의 히스토그램 공간 분포 차분에 대한 2×m 구간의 합을 R_{STotal}, G_{STotal}, B_{STotal}로서 구한다. 히스토그램 공간 합이 문턱치 이상이면 장면변화가 나타난 것으로 결정한다. 첫 번째 대표 프레임 검출 조건과 마찬가지로 동일한 문턱치로서 검출한다. 비디오의 특성에 따라, 각 블록들은 상이하게 반응한다. 장면변화가 존재하는 조건을 만족하면, 2×m 구간에서 장면변화가 가장 크게 일어나는 프레임을 대표 프레임으로 선택한다.

3.2 PFrame 추출 알고리즘

검출된 대표 프레임들은 비디오의 구성상 2×m 구간 외에서 비슷한 대표 프레임이 추출되며, 시간상으로 떨어진 영역에서 상관성이 높은 프레임이 나타날 수 있다. 멀티미디어 데이터베이스에서는 사용자 질의에 대한 정확성과, 검색 속도가 중요한 요소로 작용한다. 검출된 대표 프레임들의 상관 관계를 고려하여 색인되는 프레임 수를 줄이는 것은 데이터베이스의 검색 속도 향상을 위해서 반드시 필요하다. 각 상관성으로 재추출되는 프레임은 PFrame으로 정의한다. 대표 프레임은 각 R · G · B 히스토그램에 대한 정보가 존재하며, 블록 b에 대한 정보도 갖고 있다. 다음으로 상관성 조사가 수행된다.

```
Initialize k.
procedure PFrameExtraction
for i = 0 to Frame-1 do begin
  for j = i+1 to Frame do begin
    if correlation ( 대표프레임_i, 대표프레임_j )
        > Threshold begin
        increase k
        PFrame_k = 대표프레임_i
    end
  end
end.
function correlation(A, B : 대표프레임) begin
```

$$R_{Total} = (R_{Hist}[A] - R_{Hist}[B])$$
$$G_{Total} = (G_{Hist}[A] - G_{Hist}[B])$$
$$B_{Total} = (B_{Hist}[A] - B_{Hist}[B])$$
$$R_{STotal} = (R_{SprHist}[A] - R_{SprHist}[B])$$
$$G_{STotal} = (G_{SprHist}[A] - G_{SprHist}[B])$$
$$B_{STotal} = (B_{SprHist}[A] - B_{SprHist}[B])$$
$$Correlation = Max(R_{Total},\ G_{Total},\ B_{Total})$$
$$Correlation(+) = Max(Correlation(-),\ R_{STotal},\ G_{STotal},\ B_{STotal})$$

end.

위의 단계로 상관성 조사후, PFrame을 추출한다. 단일한 배경의 구성을 갖는 비디오들은 시간 순서상 적은 대표 프레임들이 추출되며, PFrame 추출 후에는 색인 프레임이 보다 적게 나타난다. 그림 2는 비디오 스트림에서 대표 프레임을 추출 후, PFrame을 추출하는 과정을 나타낸다.

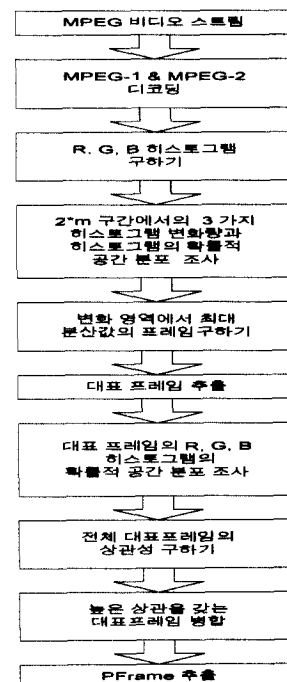



그림 2. PFrame 추출의 블록도

4. PKey 추출 알고리즘

PKey는 추출된 PFrame의 주요 키를 나타낸다. 색상 정보를 추출하기 위해 Fuzzy 합집합 연산을 이용한 양자화를 수행하여, R·G·B 색상 정보를 저장한다[9]. 방향 정보와 곡선 수를 추출하기 위해 FWSA(fuzzy worm searching algorithm)를 수행한 후, 3×3 매스크를 이용한 방향 정보와 제안하는 곡선 판별 알고리즘으로 곡선과 직선의 정보를 저장한다. 또한, 영상에 대한 기본 주성분을 PCA(principal component analysis)를 사용한다. 이러한 정보들은 하나의 집합으로 구성되어, PKey로 저장된다. 저장된 색인 키는 영상에 대한 특징을 저장하므로, MPEG-7의 서술자를 사용할 경우 간단한 PKey만으로 검색할 수 있다. 전체 영상에 대한 검색에 비해 하나의 키를 검색에 사용하므로, 수행속도의 향상과 적은 메모리를 차지하게 된다. 자세한 알고리즘은 다음과 같다.

4.1 주 색상 정보 추출

주 색상 정보는 PFrame의 가장 주요한 색상을 저장하여, 현재 키는 비디오의 색상 정보를 나타낸다. Fuzzy 합집합[10]을 이용한 양자화하는 각 히스토그램 정보를 내림차순으로 정렬하여 상위 5가지를 저장하는 경우, 뚜렷이 구별되지 않는 부분에서 비슷한 색상이 저장된다. Fuzzy 합집합은 피크에서 대표값을 취하는 방법이다. 대표값은 벡터 양자화이며 PFrame의 주요한 레벨 5개를 저장한다. R·G·B에서 5가지의 색상을 저장하고, 각 색상의 레벨은 0~255이므로, 주색상 특징점은 다음으로 계산된다.

$$size\ of\ feature[1] = 3 \times 5 \times 8\ bits$$

여기서, 3은 3가지 색상 좌표를 나타내며, 5는 주요 색상의 레벨 개수를 나타내고, 8은 0~255를 표현하는 비트 수를 나타낸다.

4.2 Fuzzy 합집합을 이용한 양자화

0과 1로 표현될 수 있는 hard set에 대응되는 개념으로, 0과 1사이에 모호한 표현의 fuzzy set 연산에 사용되는 합집합 연산을 양자화에 적용한다. 히스토그램의 레벨에 따라 오름 차순으로 정렬한 후, 구간별 퍼지 합집합 연산을 수행한다. 퍼지 합집합은 최대값을 선택한다. 선택된 레벨의 m 구간에서 최대값을 대표로 선정한다.

4.3 방향 정보 추출

방향 정보 추출은 에지 정보를 얻기 위한 전처리 과정이 필요하다. 에지 정보를 추출하는 방법으로는 Sobel, Prewitt 등의 매스크를 이용한 방법들이 있다. 본 연구에서는 퍼지의 벌레 검색 방법(FWSA; fuzzy worm searching algorithm)[11]을 이용하여 빠른 속도로 에지를 검출하여, 방향 정보와 직선·곡선 정보를 추출한다. 여기서 정의한 방향 정보는 영상 에지의 방향성을 나타내며, 하나의 특징으로 선택된다.

4.3.1 방향 마스크에 의한 정보 추출

FWSA에 의해 추출된 에지 정보로 부터 방향 정보를 저장한다. 3×3 블록 크기의 마스크를 PFrame에 사용하여, 수평(—), 수직(|), 대각선(／, ＼)의 개수를 저장한다. 다음의 마스크를 사용하여 영상의 최상단 블록에서 하단 블록으로 검색을 시작한다.

m1	m2	m3
m4	m5	m6
m7	m8	m9

m5를 기점으로 m5에 점이 존재하면 정보 추출을 시작하고, 그렇지 않은 경우 다음 블록으로 옮겨간다. m5에 점이 존재한다. (m4+m5+m6, m2+m5+m8, m1+m5+m9, m3+m5+m7)로 수직, 수평 및 대각 성분을 저장한다. 추출된 에지 정보를 특징점으로 사용할 수 있다. 각 멀티미디어 데이터베이스에서 PFrame의 방향 정보는 비디오의 성질에 따라 달라진다. 본 연구에서는 원 영상에서 손실없는 에지정보의 방향성을 얻기 위해서, FWSA를 채택하였다. 방향 정보 추출에 사용하는 매스크를 사용한 알고리즘은 블록 크기로 블록 경계 부분에 위치하는 점들은 방향성이 무시될 수 있다. 마스크 내에서의 방향성만으로도 색인과 검색에 이용할 수 있지만, 블록화 현상을 극복하기 위해 전체 영상의 에지 정보를 사용하여 직선과 곡선 정보를 추출한다.

4.4 직선 및 곡선 정보 추출

PFrame에 대한 직선과 곡선 정보를 추출하기 위해 FWSA를 에지 영상에 대하여 수행한다. 에지 성분에서 직선과 곡선을 구분할 수 있는 방법이 필요하다. 이전에 사용했던 마스크 방식을 사용하여 전 방향과 다른 방향으로 연결된 경우는 곡선으로 판단한다.

9	2	3
8	◎	4
7	6	5

위의 매스크를 이용하여 이전 점에서 현재점의 방향을 결정하게 된다. 이전의 포인트(◎)에서 현재의 방향이 좌에서 우로 이동한다면 위의 매스크에서 '4'라는 값을 가지게 되며, 위에서 아래로 이동한다면 '6'이라는 값을 얻게 된다. 이러한 과정을 수평과 수직으로 검색을 해가면서 처리해 간다. 에지 정보의 수평에 기반한 탐색을 하기 위해 수직적으로 기준점을 옮겨가면서 탐색을 수행한다. 직선과 곡선에 대한 판단은 이전 방향과 같은 경우는 직선으로 직선 개수에 포함시키고, 그렇지 않은 경우는 모두 곡선으로 판별한다. 수평 성분에 대하여 검색이 완료되면, 수직 성분에 대하여서도 탐색을 수행하고, 두 성분의 개수를 저장한다. 이렇게 추출된 직선과 곡선 개수를 특징값으로 저장한다.

4.5 PCA 정보 추출

첫 번째 특징점 추출은 히스토그램의 확률에 의해 퍼지 합집합을 이용하여 개수를 추출하였다. 주성분 정보는 빠른 양자화와 픽셀간의 오차에 따라 생성하는 주성분 해석을 통한 양자화를 이용하여 특징점을 추출한다. 일반적인 PCA에 대한 알고리즘을 사용하여 PCA정보를 추출하였다. 이들로서 PKey를 생성한다.

4.6 PKey 생성

Pkey는 4가지 특징점을 하나의 키로 저장한다. 이 키값을 색인시에 저장하고, 사용자가 검색할 경우 후보 영상을 선택하면 해당 영상의 PKey를 생성한다. 생성된 PKey와 저장된 멀티미디어 데이터베이스의 PKey들간의 유사도가 가장 높은 PFrame에 해당하는 비디오가 사용자에게 보여질 것이다. PKey 구조는 그림 4와 같다.

주요 색상	방향 성분	직선 · 곡선 개수	PCA 성분

그림 4. PKey의 구조

5. 실험 및 검토

실험 대상은 mpg, m2v 파일을 대상으로 하였으며, 윈도우 95가 탑재된 펜티엄 환경 하에서 Visual C++ 6.0으로 프로그래밍되었다. 비디오 부분만의 해석을 위하여 오디오 부분의 복호 부분은 제거하였고, 히스토그램 분석을 위해 각각의 칼라 영역으로 읽어들였다.

5.1 MPEG 비디오 복원

그림 5는 mpg 파일이 로딩된 시뮬레이션 프로그램을 나타내고 있다. 그림에서 보이는 영상은 MPEG-1 비디오의 복호기에 따라 구현된 복호 영상을 나타내고 있다.

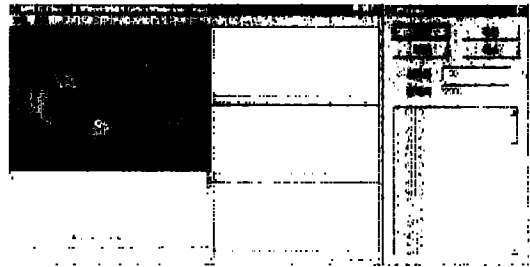



그림 5. 시뮬레이션 프로그램

그림 5에서 오른쪽 3개의 그래프는 RGB 칼라 히스토그램이다. 복원된 그림 하단은 그레이 히스토그램이다. 프로그램에 사용되는 4가지의 그래프는 모두 실시간으로 복원되는 각 프레임에서의 분석 결과이다. 분석된 결과에서 이전 프레임간의 차이가 크게 나타나는 프레임들을 추출하게 된다. 서브 프로그램으로 붙어 있는 부분에서 차이가 크게 나타난 프레임의 인덱스와 PFrame으로 추출될 것인지에 대한 것을 나타내고 있다. 서브 프로그램에서 인자로 사용되는 스케일과 문턱치는 추출된 프레임들을 PFrame으

로 추출하는데 이용하는 값이다. 실험에서는 R, G, B 각각의 색상에 대하여 100으로 스케일링하고, 현재 프레임과 이전프레임의 거리차가 2000 이하가 되는 프레임들은 변화가 적은 영역으로 판단하였다. 실험에 사용된 파일들은 기존의 통신망에서 구할 수 있는 뮤직 비디오의 MPEG파일을 이용하였다.

5.2 대표 Frame 추출 결과

뮤직 비디오와 같은 경우는 단순한 동작의 연속과 이전 장면의 뒤섞임뿐만 아니라, 복잡한 장면 변화로 구성된다. PFrame추출 이전에 추출되는 대표 프레임들은 뮤직 비디오의 특성상 많은 장면 변화로 인하여, 데이터의 수가 급증하게 된다. 이러한 프레임들에 대하여 인덱스를 부여한 후, 추출된 프레임들간의 상관성을 조사하여, 상관도가 높은 프레임들은 하나로 병합하였다. 다음 그림들은 사용한 데이터베이스에서 추출된 결과로 구성하였다. 그림 6은 뮤직비디오에서 추출된 대표 프레임들을 나타내다. 아래 그림에서 추출된 대표 프레임들은 12개 혹은 16개의 프레임을 임의로 추출하여 표현한 그림들이다.

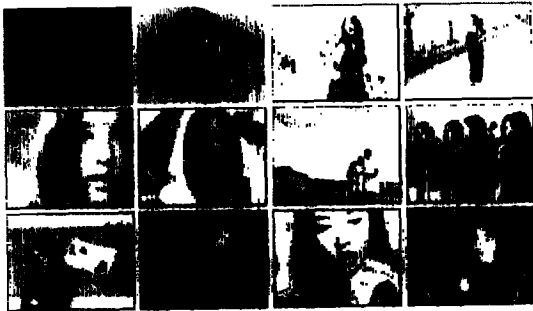



그림 6. 핑클 뮤직 비디오의 추출된 프레임

5.3 PFrame 추출 결과

추출된 대표 프레임들은 이전에 나타났던 장면들이 다시 나타나는 중복성이 존재한다. 상관성이 많은 장면들을 제거하여 PFrame을 추출한 결과가 그림 7과 같다. 그림들은 대표 프레임에서 추출된 PFrame이다. PFrame은 비디오의 장면변화에 따라서 개수가 달라진다. 표 2는 실험에 사용된 뮤직 비디오에서 추출된 대표 프레임과 PFrame의 수를 나타내고 있다.

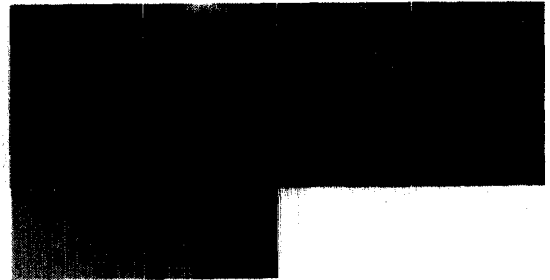



그림 7. 핑클의 뮤직비디오의 추출된 PFrame

표 2. 주요 프레임과 PFrame의 개수

뮤직 비디오	대표 프레임	PFrame
김종서 - 지금은...	82	7
동물원 - 널 사랑하겠어	73	6
수 - 웨딩마치	34	6
엄정화 - 초대	22	3
에코 - 만일 내가	187	10
조성모 - To heaven	241	28
핑클 - 내 남자 친구에게	98	10
한스 밴드 - 선생님...	40	5

5.4 PKey 추출 결과

그림 8의 시뮬레이션 프로그램에서 첫 번째 버튼은 추출된 PFrame을 선택하는 버튼이며, 두 번째 버튼부터는 "퍼지 합집합을 이용한 양자화의 주색상 정보 추출", "FWSA를 이용한 방향 정보 추출", "FWSA를 이용한 직선 및 곡선 정보 추출", "PCA 정보 추출"을 나타내고 있다. 그 아래의 "칼라", "빨강", "초록", "파랑" 버튼은 그림에 대한 칼라와 R · G · B 의 영상을 표시하고 있다. 우측 부분의 공간은 해당 PKey 추출후의 결과이다. 디스플레이 부분은 원 영상에 대한 B 색상의 영상을 표시하고 있으며, 결과를 표시하는 부분에서는 "색상[인덱스 : 레벨] = 색상수"를 나타내고 있다. 색상수는 3가지 레벨의 색상이 가지는 수이다. 그림 에서는 퍼지 벌레 검색 알고리즘을 이용한 방향 정보 추출 결과를 보여주고 있다. 왼쪽 디스플레이 부분은 FWSA를 적용한 R 색상값을 보여 주고 있다. 매스크를 사용한 에지 검출에 비하여 원 영상에 대한 무손실의 추출 영상이다. 결과 출력 부분에서 표시하고 있는 의미는 "색상[방향 인덱스] = 방향 결과 수치"와 같다. 방향 인덱

스에서 0은 ―, 1은 \, 2는 |, 3은 ╱의 방향을 나타낸다. 방향 결과 수치는 앞의 알고리즘에서 설명한 방향 매스크를 사용하여 합산된 결과 수치를 표시한다.

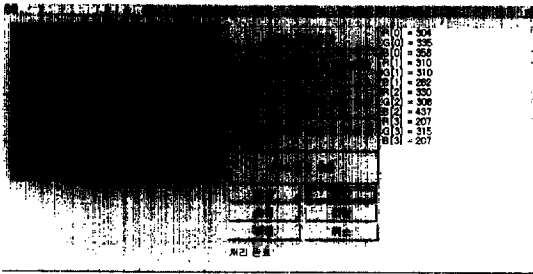



그림 8. FWSA에 기반한 방향 매스크를 사용한 방향 정보 추출

그림 9는 FWSA를 이용한 직선 및 곡선 정보를 추출후의 PFrame에서 주성분 분석을 이용하여 주성분을 추출한 결과를 나타낸다. 왼쪽의 PFrame에서 G 색상에 대한 정보를 나타내며, 오른쪽은 3가지 색상에 대한 주성분의 추출 결과이다. 결과의 x와 y는 PFrame 에서의 홀수번째 픽셀들과 짝수번 픽셀을 매핑한 값을 나타내고 있다. 주성분 추출은 매핑된 값에 대한 주축을 구한 뒤에 확률 분포에 따라서 주성분에 직각인 법선 벡터를 구하고, 구해진 법선 벡터에 의하여 두 부분으로 나뉜다. 나누어진 부분에 대하여 다시 주축을 구하고 구해진 주축에서 확률에 따른 법선을 구하는 과정을 원하는 노드수 만큼 반복하게 된다. 결론적으로, x와 y값을 가지게 되므로 각 색상당 총 8가지의 성분을 저장하게 된다.

그림 9. 추출된 PCA 정보

이상과 같이 추출된 정보는 팩(pack)되어 하나의 PKey로서 저장된다. 3~5분에 이르는 뮤직 비디오들에 대한 개별적인 패턴 매칭 방법은 "전체 프레임 수 × 픽셀 수 × 영상 크기"로 표현될 수 있으나, PKey에 의한 매칭은 PFrame에서 추출된 정보이므로 "PKey 특징 개수 × PKey 개수"의 검색 시간을 갖는다. 따라서, 기존의 후보 영상들로서 비교하는 타 알고리즘보다 수행속도를 높일 수 있다.

6. 결 론

본 논문에서는 MPEG-7의 기반 기술로서 비구문 기반 서술자를 생성하기 위한 PFrame을 정의하고 생성하는 방법을 제안하였으며, PFrame에서 색인과 검색을 위한 PKey를 추출하는 방법에 대하여 제안하였다. 인간의 지각적인 능력에 기인하여, 색상과 객체의 변화에 민감한 특성을 고려한 색인에 적용한다. 구문 기반 서술자에 비하여 일관적인 특성으로 멀티미디어 비디오를 색인하므로서, 검색 시스템의 효율을 높일 수 있으며, 사용자와 개발자간의 의사소통이 필요없으므로 정확한 검색 시스템을 구성할 수 있다. 실제 실험에서 사용된 뮤직비디오에 대한 PKey들은 최소 검색키로서 수행 속도를 향상시켜며, 사용자의 검색에 대한 신뢰성을 높일수 있었다. 앞으로의 연구 과제는 비디오뿐만 아니라 오디오 특성에 따른 특징점을 추출하여 통합된 색인과 검색 시스템을 구성하는 것이다.

참 고 문 헌

[1] M. Sezan, R. Qian, "MPEG-7 Standardization Activities", Proceedings IEEE international conference on image processing, Vol. 3, pp. 517-520, 1998.

[2] M. Naphade, T. Kristjansson, B. Frey, T. Huang, "Probabalistic Multimedia Objects (Multijects): A Novel Approach to Video Indexing and Retrieval in Multimedia Systems", Proceedings IEEE international conference on image processing, Vol. 3, pp. 536-540, 1998.

[3] Chueh-Wei Chang, Suh-Yin Lee, "Video Content Representation, Indexing, and Matching in Video Information Systems", Journal of Visual Communication and Image Representation, Vol. 8 No. 2, pp. 107-120, 1997.

[4] J. Nam, M. Alghoniemy, A. Tewfik, "Audio-Visual Content-based Violent Scene Characterization", Proceedings IEEE international conference on image processing, Vol. 1, pp. 353-357, 1998.

[5] 서병찬, 이동호, "MPEG-2 비디오 디코더 구현을 위한 Variable Length Decode 설계", 제11회 신호처리종합학술대회 논문집, Vol. 11 No. 1, pp 179-182, 1998.

[6] J. S. Park and J. H. Han, "Contour motion estimation from image sequences using curvature information", Pattern recognition, Vol. 31 No. 1, pp. 31-40, 1998.

[7] N. Doulamis, A. Doulamis, Y. Avrithis, Y. Aoki, "Indexing of Baseball Telecast for Content-based Video Retrieval", ibid., Vol. 1, pp. 875-879, 1998.

[8] H. Xu, M. Liao, "Cluster-Based Texture Matching for Image Retrieval", ibid., Vol. 2, pp. 766-769, 1998.

[9] A. Pinho, "Encoding of Closed Boundaries using Transition Points", ibid., Vol. 1, pp. 271-275, 1998.

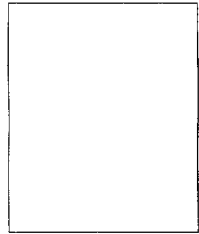
[10] Hou-Chun Ting, Hsueh-Ming Hang, "Edge Preserving Interpolation of Digital Images Using Fuzzy inference", Journal of Visual Communication and Image Representation, Vol. 8 No. 4, pp 338-355, 1997.

[11] Seong Wook Ha, Serk Bae Ser, Dae-Seong Kang, "Image Segmentation Based on Combination of the Genetic Fuzzy Worm Searching and Genetic Adaptive Min-Max Clustering", ITC-CSCC' 98, Vol. 1, pp 201-204, 1998.

하 성 욱

1997년 동서대학교 컴퓨터공학과 졸업(학사)
1999년 동아대학교 전자공학과 졸업(석사)
1999년~현재 동아대학교 컴퓨터공학과(박사과정 재학중)
관심분야 : MPEG, 악보 인식, 청음
e-mail : g9773111@seunghak.donga.ac.kr

강 대 성

1984년 경북대학교 전자공학과(학사)
1991년 Texas A&M Univ. Electrical Eng.(석사)
1994년 Texas A&M Univ. Electrical Eng.(박사)
1984년~1989년 국방과학연구소 연구원
1994년~1995년 한국전자통신연구원 선임연구원
1995년~현재 동아대학교 전기전자컴퓨터공학부 조교수
관심분야 : 영상처리, 영상압축, 멀티미디어통신, 패턴인식 등
e-mail : dskang@daunet.donga.ac.kr

김 대 진

1981년 연세대학교 전자공학과(학사)
1984년 KAIST 전기 및 전자공학과(석사)
1991년 Syracuse 대학 ECE(박사)
1984년~1987년 KBS 기술연구소 HDT팀
1992년~1999년 동아대학교 컴퓨터공학과 부교수
1999년~현재 포항공과대학교 컴퓨터공학과 부교수
e-mail : dkim@postech.ac.kr

권 기 항

1983년 서울대 컴퓨터공학과 졸업(학사)
1985년 미국 Georgia Tech 컴퓨터과학과 졸업(석사)
1986년~1988년 ㈜상운
1995년 미국 Duke 대학 컴퓨터과학과 졸업(박사)
1995년~현재 동아대학교 컴퓨터공학과 조교수
관심분야 : 소프트웨어공학, 프로그래밍 언어
e-mail : khkwon@daunet.donga.ac.kr